

# Machine Encoded Text Based Product Identification for Visually Challenged

A.Dinesh Babu<sup>1</sup>, B.Dheeraj<sup>2</sup>, J.Jose Samuelraj<sup>3</sup>

Assistant Professor<sup>1</sup>, UG Scholars<sup>2,3</sup>

dbanbumani@gmail.com<sup>1</sup>, b.d.shailesh@gmail.com<sup>2</sup>, josesamuel96@gmail.com<sup>3</sup>

Department of Electronics and Communication Engineering,  
SRM Institute of Science and Technology, Vadapalani, Chennai-600 026, India

**Abstract**--Technology is insufficient to provide facility to the visually challenged for identification of products. Thus, the major aim of this work is to help the visually challenged for identifying products. Text and character identification by machine encoding is way to help the visually challenged for product identification. By machine encoding the text, the characters and text are recognized and then it is synthesized to give voice output. This encoding is done by using Raspberry Pi with the Open Cv libraries. In this work, the first step is to consider all the printed or handwritten characters as inputs. Our proposed work is helpful in automatic detection of the objects with the help of Ultrasonic sensor, and also a text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels using Open Cv libraries is offered. Text characters in the localized text regions are then converted into binary bits later recognized by off-the-shelf optical character identification software. The renowned text codes are converted into audio outputs for the visually challenged. Usage of Ultrasonic sensor in this work is the major modification of product identification. The object distance is measured by ultrasonic sensor which automatically focuses the text regions from the object. Thus, we overcome the difficulty to identify the position of the object for visually challenged. Visually challenged people can use this for product identification and it is highly beneficial for them.

**Keywords** —Image recognition; Ultrasonic Sensor; Text extraction; Product Identification.

## I.INTRODUCTION

Machine encoded text is the electronic conversion of pictures by providing alphanumeric recognition of printed or handwritten characters. OCR has been an active topic of research in the recent past, and has wide applications in banking, healthcare, finance and education. OCR is a field of research in Computer Vision, Artificial Intelligence and Pattern Recognition [1]. Thus there is a pressing need to develop a system which is cheaper in cost and affordable to the needed. This work addresses a complete machine which converts the text to audio system built on an embedded framework. The proposed idea uses the principle of integrating a camera based assistive device and Raspberry Pi3 board. The integrated system consists of a camera module, webcam, Tesseract Optical Character Recognition Engine, Pico Text to Speech Engine, Speakers/Headphones and Computer Vision software for picture processing. The camera is an input device which feeds the required pictures for digitization. These pictures are processed using OpenCV libraries. The processed pictures serve as an input to the OCR, which digitizes the pictures and performs character recognition. Finally, the Text to Speech engine reads out the text to the user. It is an end to end trainable network for text identification [2]. The steps, proposed system design and experimental outcomes are discussed in this paper work.

## II.LITERATURE SURVEY

Machine encoded text has been a dynamic subject of research since 10 years. The quick development of advanced libraries overall postures new difficulties for archive picture investigation innovative work. A dream based content acknowledgment framework utilizing Raspberry pi is displayed in [3]. This paper proposes a total optical acknowledgment framework. The picture is first caught using the Raspberry Pi camera module or using a webcam. Picture pre-processing systems are connected where the undesirable commotion is expelled by applying limit morphological changes - morphological, widening, dark cap and discrete cosine. Content regions are removed by drawing jumping boxes around the required content and the inclination is adjusted. The flat and vertical proportions are changed in accordance with dispose of undesirable high recurrence segments. Subsequent to thresholding, forms of the picture are created by utilizing uncommon OpenCV capacities.

The consequence of picture to content transformation is passed on to a discourse motor, which is fit for changing over content to discourse utilizing specific libraries. The paper utilizes TTS Festival for this reason. A keen peruse for the outwardly weakened utilizing Raspberry Pi is talked about in [4]. The reproduction of optical character acknowledgment is finished utilizing Mat lab. The caught picture is sent for pre-handling where capacities, for example, skew redress, commotion evacuation and linearization are performed. The picture is first lit up and afterward binarized. The picture is then passed on to the division stage where the picture is decayed into characters. This picture is then sent to the TTS motor. An optical character age framework for change of pictures to content is displayed in [5]. The parsing comprises of three stages: Character extraction, Recognition and post preparing. Amid the acknowledgment stage, the format with most extreme connection is pronounced as the character show in the picture.

A Neural Network (NN) based procedure for optical character acknowledgment and written by hand character acknowledgment is exhibited in [6]. This paper additionally talks about different procedures like framework coordinating, fluffy rationale, include extraction and basic investigation. The neural system show utilized here is the Multi-Level Perception demonstrate (managed organize). The yields are given as a piece of the preparation vector. The last yield of the system is dictated by actuations from yield layer. Another calculation for optical character acknowledgment which exploits the typographic consistency of sections or other format segments is proposed in [7]. The quintessence of versatile Machine encoded text is to lessen multi-text style characterization by exploiting the ordinary event of long series of characters in a similar typeface. Typeface homogeneity likewise helps increment the throughput. A framework that perceives printed content of different textual styles and sizes for the Roman letter set is proposed in [8]. Shape extraction is performed specifically on the diagram of the run-length encoding of a twofold picture utilizing a shape bunching approach.

This is then nourished into a Bayesian Classifier. At long last, design and etymological setting are connected. A system where the double pixels are prepared utilizing a syntactic approach that marks each sensibly noteworthy square to the page as indicated by the pre-aggregated data is proposed in [9]. Machine encoded text is utilized to change over content pieces to ASCII frames. Picture and content squares are put away in X-Y tree information structure that can be gotten to through a LAN or WAN.

A finger worn gadget to help the outwardly debilitated with perusing printed books is displayed in [10]. It extricates content from a nearby camera see which empowers ceaseless criticism. An implanted optical character acknowledgment on Tamil content pictures utilizing Raspberry Pi is proposed in [11]. The Tamil content picture is checked and separated, and is pre-processed. Pre-handling comprises of character division also, picture linearization. The highlights are extricated and the picture is sent for post-preparing which changes over the picture to content. This content is then encouraged to a discourse motor which produces blended Tamil discourse as yield. comprises of a multimodal criticism by means of vibration engines, another double material case outline and a high-determination smaller than expected camcorder. Tesseract Machine encoded text and Flite Text to Speech has been utilized. It is executed in a novel following calculation that

An overview of techniques and methodologies in character division is talked about in [12]. The customary approach segments the picture into sub pictures which are then arranged autonomously. The below average of division sections the picture either unequivocally or verifiably through the spatial highlights gathered. The second rate class is in the middle of the initial two that utilizes dismemberment alongside recombination.

### III. MACHINE ENCODING PROCEDURE

In this process of text identification, three steps are part of it they are, segmentation, Pictures Extraction, and Description. From this paperwork we tell how our steps deal with all hindrances. At division the substance of a pictures is settled, it is important that the pictures and movements are should have been settled, characters are distinguished independently and are isolated, this progression is simple yet challenges happen when characters consolidate with each. Loud pictures and shapes in the pictures are few inconveniences in the picturesin addition because of the consolidated characters and appended writings with movements, the recognizable proof advance doesn't get the contribution for distinguishing proof.

The subsequent stage is getting to the primary highlights of the considerable number of characters which separate every one of the images, it is the most muddled work to be done in character recognizable proof, include extraction is done demurely by inspecting the allotment of focuses, by transformation, auxiliary examination and arrangement addition and few highlights genealogy techniques are format closeness, changeover, assignment of speck by tract, trademark locus, n-tuple, journey and by basic examination. The following stage of distinguishing proof is the division, analysing all the character and apportion to the correct character sets. The two courses by which division should be possible are the choice theoretic route, spoke to by numbers byvector. Another way is the basic technique, it is done when there is association among highlights of characters by cases in the event that we realize that the character contains of onelatitude line and one longitudinal queue, so we can expect that the character can be 'L' or 'T'. In this printed material we have relegated the broadening byneural system by the assistance of back proliferation method. The method for ID of writings relies upon pictures determination i.e., pixels. The reviewed

pictures are mind boggling due to changes in foundation and writings. In this printed material we will examine the condition sets that work on various kinds of pictures with different text dimensions, messages and hues.

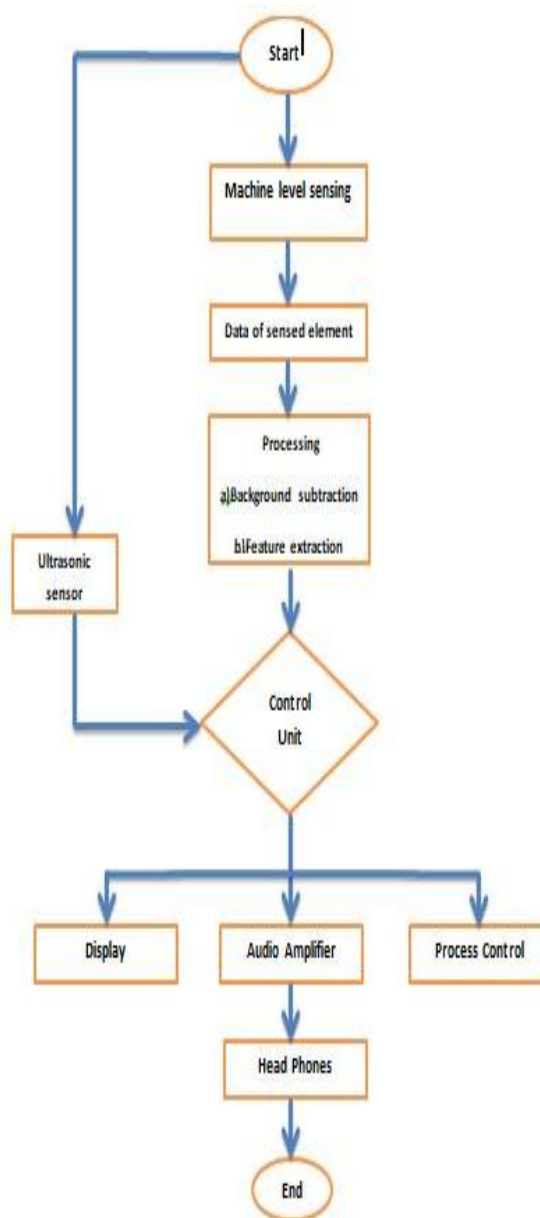


Figure 1. DesignFlow of Encoding Text

The process flow is shown in the above flowchart Figure 1 where the steps of the machine encoding text process takes place. Feature extracted and background subtracted images are character identified and the input is given to the control unit and the voice output is received.

### IV. PROPOSED SYSTEM DESIGN

The pictures are scanned using the Web Camera. It is pre-processed by Background subtraction and feature lineage. The Power unit is directly connected to the control unit. The control unit we have used here is the Raspberry pi. It complies of the Open Cv libraries and the machine encoding texts. The control unit extracts the texts and the Text-to-Speech converter works by which we receive the audio output. The block diagram is shown in Figure 2.

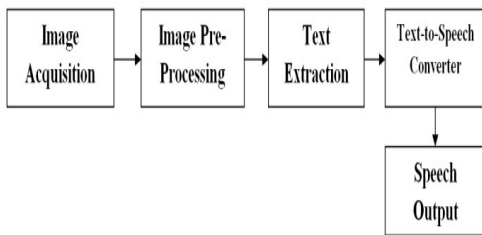


Figure 2 Block Diagram

The pictures of a printed text is captured through the Raspberry Pi camera module[4]. This image is subjected to pre-processing which includes correcting skew angles, sharpening of pictures, thresholding and segmentation. The processed image is sent to the TTS synthesizer. The architecture of this design is given in Figure 3.

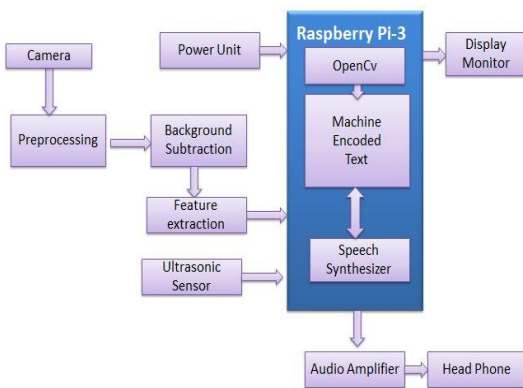


Figure 3 System Architecture

The picture is scanned and the scanned image is background subtracted and their features are extracted. This processed image is now read by the control unit, The control unit Raspberry pi is connected with Keyboard, Mouse, Display, Power adapter(5V) and Web Camera. The captured picture is processed and it is scanned by the tesseract by the control unit. The speech output is received from the control board using earphones.

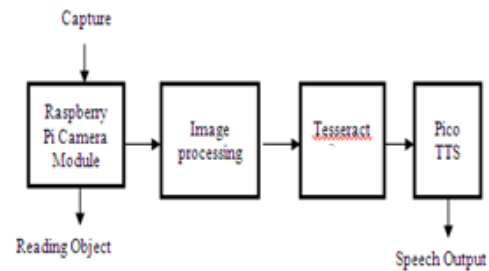


Figure 4 System design

### V. EXPERIMENTAL OUTCOME

This method has been examined on fractioned characters, dull characters, spread characters & also on pictures which consists graphics and combined characters. It is examined that in before processing of pictures that we use various methods to delete unnecessary graphics, sound and unnecessary text, the appropriate character identification is incremented. When hamming distance and correlation is used by feature lineage as frame similarity method with the operation of back propagation method as division we get the maximum output for machine encoded text, it is applied on various pictures by changing the pictures size, such as the pictures of various pixel. It has been examined on low quality pictures i.e., 4x4.

Pictures can be chopped, resized, turned right or left, zoomed in zoomed out, and the resolution of pictures can be changed, if unnecessary text, will be deleted from pictures more noise in the sample pictures will be removed, our work shows that if pre-processing of pictures is used exactly the appropriate character identification is developed at greater level. All the modules of pre-processing is developed by using Python programming language. The Open CV library has inbuilt characters to be identified. The scanned pictures are read by using the libraries to identify the accent of the characters. Identification of the text by the library is an inbuilt process

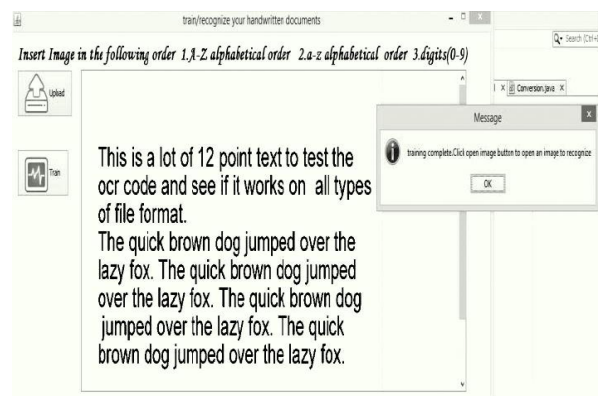


Figure 5 Development of character Set

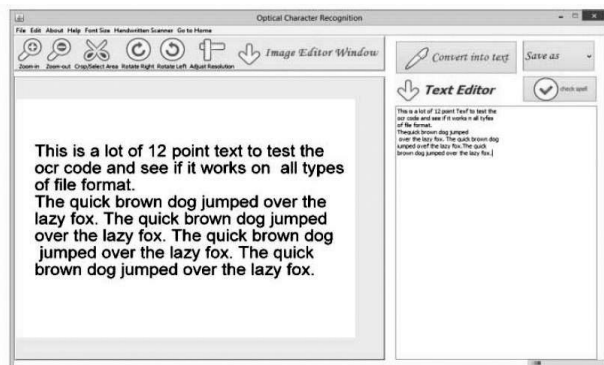


Figure.6. Text Identification

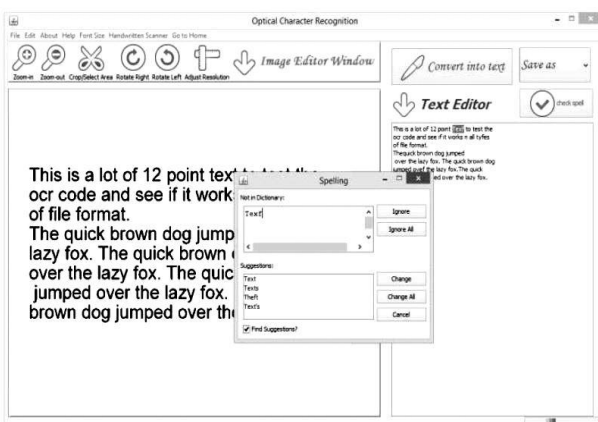


Figure 7. Post Processing Through Open CV

### VI. CONCLUSION

In this paperwork we encased whole technique to utilize character recognizable proof, it has been inspected that layout comparability element heredity way is executed, we analysed that element ancestry is for the most part the primary issue for getting suitably. Text and character identification by machine encoding is way to help the visually challenged for product identification. By machine encoding the text, the characters and text are recognized and then it is synthesized to give voice output. This encoding is done by using Raspberry Pi with the Open Cv libraries. The objects distance is identified with the help of Ultrasonic sensor which automatically detects the products to visually challenged for identification. Nearly the pre-preparing routes encased in this printed material likewise decrease clamour at a more prominent level. The printed material examines about the technique connected in character ID utilizes create utilizing Python. Optical characters can be recognized in various techniques, although executing something that imitates the nature and highlight of human mind will undoubtedly convey superior to some other procedure. The proposed framework is executed by Open Cv Library for crafted by machine encoded content, this work is highly beneficial for visually challenged. Visually challenged people can use this for project identification. In future, there is go for such a procedure to use in robot and in addition for accord human composed substance. Such an assignment can be additionally expanded with the end goal that it can be used to peruse other languages.

### VII. REFERENCES

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton “Pictures net classification with deep convolutional neural networks,” in Proc. Advances in Neural Info. Processing Sys., 2012, pp. 1106–1114.
- [2] K. Wang, B. Babenko, and S. Belongie, “End-to-end scene text recognition,” in Proc. Int. Conf. Comput. Vision, 2011, pp. 1457–1464.
- [3] Nagaraja L. et al, “Vision Based Text Recognition using Raspberry Pi”, International Journal of Computer Applications, NCPSIA 2015.
- [4] Velmurugan D. et al, “Hardware Implementation of Smart Reader for Visually Impaired using Raspberry Pi” ,IJAREEIE, Vol.5 , Issue 3, March 2016, pp 2055-2063.
- [5] Abhishek Verma et al, “OCR- Optical Character Recognition”, IJARSE, Vol. 5, Issue No 09, September 2016, pp 181-191.
- [6] N. Venkata Rao et al, “Optical Character Technique Recognition Algorithms, Journal of Theoretical and Applied Information Technology”, Vol.83, No. 2, 20th January 2016.
- [7] George Naggy, “At the Frontiers of OCR”, Invited Paper, pp 1093-1100.
- [8] Richard G. Casey et al, “A survey of methods and strategies in Character Segmentation”, IEEE Computer Society, July 1996.
- [9] George Naggy et al, “A Prototype Document Image Analysis System for Technical Journals”, University of Nebraska - Lincoln Digital Commons, at University of Nebraska – Lincoln, pp 10-22.
- [10] Roy Shilkrot et al, Finger Reader: “A Wearable Device to Support Text Reading on the Go”, MIT Research Paper.
- [11] Capt Dr S Santhosh Baboo et al, “Embedded Optical Braille Recognition on Tamil Braille System using Raspberry Pi”, IJCTA, Vol.5, pp 1566-1574.
- [12] K. Nirmala Kumari et al, “Image Text to Speech Conversion using OCR Technique in Raspberry Pi”, IJAREEIE, Vol 5, Issue 5, May 2016.



